https://indico.jinr.ru/event/1086/contributions/

9th International Conference "Distributed Computing and Grid Technologies in Science and Education" (GRID'2021)

Список докладов

196. JINR's strategic plan for long-term development

Dr Grigory Trubnikov (JINR)

190. Information Technologies @ JINR development strategy

Vladimir Korenkov (JINR)

202. System programming and cybersecurity

Arutyun Avetisyan (ISP RAS)

187. Distributed scientific computing challenges and outlook

Oxana Smirnova (Lund University)

60. Can I protect my face image from recognition?

Nadezhda Shchegoleva (Saint Petersburg Electrotechnical University "LETI")

05.07.2021, 15:30

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The "Fawkes" procedure is discussed as a method of protection against unauthorized use and recognition of facial images from social networks. As an example, the results of an experiment are given, confirming the fact of a low result of face image recognition within CNN, when the "Fawkes" procedure is applied with the parameter mode = "high". Based on a comparative analysis with the original...

74. IHEP tier-2 computing center: status and operation

Viktor Kotliar (IHEP)

Distributed computing systems

RU-Protvino-IHEP site is the one of three biggest WLCG Tier-2 centers in Russia. The computing infrastructure serves for "big four" LHC high energy physics experiments such as Atlas, Alice, CMS, LHCb and local experiments at IHEP such as OKA, BEC, radio biology stands and others. In this work the current status of the computing capacities, networking and engineering infrastructure is shown as...

35. The ATLAS EventIndex using the HBase/Phoenix storage solution

Elizaveta Cherepanova (Laboratory of Nuclear Problems)

3. Computing for MegaScience Projects

Computing for MegaScience Projects

The ATLAS EventIndex provides a global event catalogue and event-level metadata for ATLAS analysis groups and users. The LHC Run 3, starting in 2022, will see increased data-taking and simulation production rates, with which the current infrastructure would still cope but may be stretched to its limits by the end of Run 3. This talk describes the implementation of a new core storage service...

97. Программный интерфейс для функционального программирования для параллельных и распределенных систем

Ivan Petriakov (Saint Petersburg State University)

Sectional reports

Distributed computing applications

Существует огромное количество научных и коммерческих приложений, написанных с прицелом на последовательное исполнение. Запуск таких программ на многопроцессорных системах возможен, но без использования преимуществ этих систем. Для выполнения программы с учетом этих возможностей зачастую необходимо переписать программу. Однако, это не всегда оптимальный выбор. В этой работе рассматривается...

54. An intelligent environmental monitoring platform

Alexander Uzhinskiy (Dr.)

4. Distributed computing applications

Sectional reports

Distributed computing applications

Air pollution has a significant impact on human and environmental health. The aim of the UNECE International Cooperative Program (ICP) Vegetation in the framework of the United Nations Convention on Long-Range Transboundary Air Pollution (CLRTAP) is to identify the main polluted areas of Europe, produce regional maps and further develop the understanding of the long-range transboundary...

55. INP BSU grid site

Dmitry Yermak (Institute for Nuclear Problems of Belarusian State University)

1. Distributed computing systems

Sectional reports

Current status of the INP BSU grid site. An overview of INP BSU computational facilities usage and cloud resources integration with JINR cloud is presented.

24. On methods of the transfer learning in the classification of the biomedical images

Проф. EUGENE SHCHETININ (Financial University under the Government of Russian the Federation)

9. Big data Analytics and Machine learning

Big data Analytics and Machine learning.

In this paper, computer studies of the effectiveness of the use of transfer learning methods for solving the problem of recognizing human brain tumors based on its MRI images are carried out. The deep convolutional networks VGG-16, ResNet-50, Inception_v3, and MobileNet_v2 were used as the basic models. Based on them, various strategies for training and fine-tuning models for recognizing brain...

37. Performance testing framework for the ATLAS EventIndex

Elizaveta Cherepanova (Laboratory of Nuclear Problems)

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The ATLAS EventIndex is going to be upgraded in advance of LHC Run 3. A framework for testing the performance of both the existing system and the new system has been developed. It generates various queries (event lookup, trigger searches, etc.) on sets of the EventIndex data and measures the response times. Studies of the response time dependence on the amount of requested data, and data...

49. Deep Learning Application for Image Enhancement

Ahmed Elaraby (South Valley University, Egypt)

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Recently, deep learning has obtained a central position toward our daily life automation and delivered considerable improvements as compared to traditional algorithms of machine learning. Enhancing of image quality is a fundamental image processing task and. A high-quality image is always expected in several tasks of vision, and degradations like noise, blur, and low-resolution, are required...

36. Development of the ATLAS Event Picking Server

Evgeny Alexandrov (JINR)

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

During LHC Run 2, the ATLAS experiment collected almost 20 billion real data events and produced about three times more simulated events. During physics analysis it is often necessary to retrieve one or a few events to inspect their properties in detail and check their reconstruction parameters. Occasionally it is also necessary to select larger samples of events in RAW format to reconstruct...

30. Research Cloud Computing Ecosystem in Armenia

Mr Artashes Mirzoyan (Institute for Informatics and Automation Problems of the National Academy of Sciences of the Republic of Armenia)

1. Distributed computing systems

Sectional reports

Research Cloud Computing Ecosystem in Armenia

Abstract
Growing needs for computational resources, data storage within higher-educational institutions and the requirement for a lot of investment and financial resources the idea or the concept of “National Research Cloud Platform (NRCP)” is crucial to provide necessary IT support for educational, research and development activities,...

28. The technology and Tools for the Building of Information Exchange Package Based on Semantic Domain Model

Elena Yasinovskaya (Plekhanov Russian University of Economics)

6. Data Management, Organisation and Access

Sectional reports

Distributed computing applications

This paper presents the technology developed by the authors to improve the semantic interoperability of heterogeneous systems exchanging information through an object-oriented bus. We demonstrate the solution that allows semantically map the models of interacting information systems with a unified data model (domain otology) when developing an information exchange package.

99. Architecture of a generative adversarial network and preparation of input data for modeling gamma event images for the TAIGA-IACT experiment

Yulia Dubenskaya (SINP MSU)

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Very-high-energy gamma ray photons interact with the atmosphere to give rise to cascades of secondary particles - Extensive Air Showers (EASs), which in turn generate very short flashes of Cherenkov radiation. This flashes are detected on the ground with Imaging Air Cherenkov Telescopes (IACTs). In the TAIGA project, in addition to images directly detected and recorded by the experimental...

89. COMPASS production system: Frontera experience

Artem Petrosyan (JINR)

1. Distributed computing systems

Sectional reports

Distributed computing systems

Since 2019, the COMPASS experiment works on the Frontera high performance computer. This is a large machine (number 5 in the ranking of the most powerful supercomputers in 2019) and details, problems, and approaches to organizing data processing on this machine are presented in this report.

100. Computing environment for the Super-Charm-Tau factory detector project

Dmitriy Maximov (Budker Institute of Nuclear Physics)

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The project of the Super Charm-Tau (SCT) factory --- a high-luminosity electron-positron collider for studying charmed hadrons and tau lepton

--- is proposed by Budker INP. The project implies single collision point equipped with a universal particle detector. The Aurora software framework has been developed for the SCT detector. It is based on trusted and widely used in high energy...

34. Development of dashboards for the workflow management system in the ATLAS experiment

Aleksandr Alekseev (National Research Tomsk Polytechnic University)

4. Distributed computing applications

Sectional reports

Distributed computing applications

The UMA software stack developed by the CERN-IT Monit group provides the main repository of monitoring dashboards. The adaptation of this stack to the ATLAS experiment began in 2018 to replace the old monitoring system. Since then, many improvements and fixes have been implemented to the UMA. One of the most considerable enhancements was the migration of the storage for aggregated data from...

128. Concurrently employing resources of several supercomputers with ParaSCIP solver by Everest platform

Sergey Smirnov (Institute for Information Transmission Problems of the Russian Academy of Sciences)

1. Distributed computing systems

Sectional reports

Distributed computing systems

ParaSCIP is one of the few open-source solvers implementing a parallel version of the Branch-and-Bound (BNB) algorithm for discrete and global optimization problems adapted for computing systems with distributed memory, e.g. for clusters and supercomputers. As is known from publications there were successful using up to 80,000 CPU cores during solving problems from the MIPLIB test libraries....

39. Lifecycle Management Service for the compute nodes of Tier1, Tier2 sites (JINR)

Alexandr Baranov ((JINR))

2. Research infrastructure

Sectional reports

Distributed computing applications

Megascience experiments, such as CMS, ATLAS, ALICE, MPD, BM@N, etc., are served at the Meshcheryakov Laboratory of Information Technologies (MLIT) of the Joint Institute for Nuclear Research (JINR) using the

available computing infrastructure. To ensure the guaranteed and stable operation of the infrastructure under constant load conditions, the centralized and timely maintenance of software...

120. Simulation Model of an HPC System for Super Charm-Tau Factory

Dmitry Wiens (ICMMG SB RAS)

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

This work describes the design of a digital model of an HPC system for processing data from the Super Charm-Tau factory electron-positron collider of the "megascience" class. This model is developed using the AGNES multiagent modeling platform. The model includes intelligent agents that mimic the behavior of the main subsystems of the supercomputer, such as a task scheduler, computing...

179. The use of convolutional neural networks for processing stereoscopic IACT images in the TAIGA experiment

Stanislav Polyakov (SINP MSU)

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Machine learning methods including convolutional neural networks (CNNs) have been successfully applied to the analysis of extensive air shower images from imaging atmospheric Cherenkov telescopes (IACTs). In the case of the TAIGA experiment, we previously demonstrated that both quality of selection of gamma ray events and accuracy of estimates of the gamma ray energy by CNNs are good...

219. Machine Learning for Data Quality Monitoring at CMS Experiment

Ilya Gorbunov (JINR)

Sectional reports

Big data Analytics and Machine learning.

We give an overview of the CMS experiment activities to apply Machine Learning (ML) techniques to Data Quality Monitoring (DQM).

In the talk special attention will be paid to ML for Muon System and muon physics object DQM. ML application for data certification (anomaly detection) and release validation will be discussed.

185. Participation of Russian institutes in the processing and storage of ALICE data

Andrey Zarochentsev (SPbSU)

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The report presents the results of the work of Russian institutes in the processing of ALICE experiment data during the last 3 years of the operation of the Large Hadron Collider (LHC) including the end of the LHC RUN2 and the1st year of the COVID-19 pandemic. The main problems and tasks facing both ALICE Grid Computing and its Russian segment before the LHC RUN3, including the problems of...

96. WALT Platform for Web Application Development

Ivan Sokolov (Alexandrovich)

1. Distributed computing systems

Sectional reports

Distributed computing applications

At the moment, there are many different platforms for web-applications developing: Django, ASP.NET Core, Express, Angular, etc. Usually, these platforms assume a division of labour when a relatively large group of developers are working on a project, each of whom is engaged in its part
(design, layout, front-end, back-end).
In our real life, usually, only 1-2 people (full-stack developers)...

48. Опыт организации гибкого доступа к удаленным вычислительным ресурсам из среды JupyterLab с использованием технологий проектов Everest и Templet

Dr Sergey Vostokin (Samara National Research University)

1. Distributed computing systems

Sectional reports

Развитие технологий искусственного интеллекта и больших данных (big data) явилось стимулом разработки новых инструментальных средств организации и автоматизации рабочих процессов (workflow). Проект Jupyter – один из основных проектов автоматизации рабочих процессов в области искусственного интеллекта. Ключевыми парадигмами проекта являются клиент-серверная модель и графическая интерактивная...

198. Evolution of the WLCG computing infrastructure for the High Luminosity challenge of the LHC at CERN

Simone Campana (CERN)

Plenary reports

Plenary reports

205. Perspective and Strategy of IT Development at IHEP

Qiulan Huang (Institute of High Energy Physics, CAS, China)

06.07.2021, 09:40

Plenary reports

Plenary reports

177. Status of the DIRAC Interware Project

Andrei Tsaregorodtsev (CPPM-IN2P3-CNRS)

06.07.2021, 10:30

1. Distributed computing systems

Plenary reports

Plenary reports

DIRAC Interware is a development framework and a set of ready to use components that allow to build distributed

computing systems of any complexity. Services based on the DIRAC Interware are used by several large scientific

collaborations such as LHCb, CTA and others. Multi-community DIRAC services are also provided by a number of

grid infrastructure projects, for example EGI, GridPP...

110. dCache: Inter-disciplinary storage system

Tigran Mkrtchyan (DESY)

06.07.2021, 11:15

6. Data Management, Organisation and Access

Plenary reports

Plenary reports

The dCache project provides open-source software deployed internationally to satisfy ever more demanding storage requirements. Its multifaceted approach provides an integrated way of supporting different use-cases with the same storage, from high throughput data ingest, data sharing over wide area networks, efficient access from HPC clusters and long term data persistence on a tertiary...

218. Intel architecture, technology and products for HPC and GRID. How to create the most effective system.

Nikolay Mester (INTEL)

06.07.2021, 12:00

Plenary reports

217. Disaggregated infrastructure: future trend and current implementation in MLIT JINR Govorun system.

Alexander Moskovsky

06.07.2021, 12:20

Plenary reports

Plenary reports

69. Concept of peer-to-peer caching database for transaction history storage as an alternative to blockchain in digital economy

Mikhail Belov (Dubna State Univeristy)

06.07.2021, 13:30

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

The development of the digital economy implies storing the history of a large number of transactions of every citizen involved in business processes based on digital technologies, starting from receiving public and social services in electronic form and ending with consumption of electronic goods and services produced by e-business and e-commerce.
If we look carefully at the data structure...

88. Event index based correlation analysis for the JUNO experiment

Tao Lin (IHEP)

06.07.2021, 13:30

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The Jiangmen Underground Neutrino Observatory (JUNO) experiment is mainly designed to determine the neutrino mass hierarchy and precisely measure oscillation parameters by detecting reactor anti-neutrinos. The total event rate from DAQ is about 1 kHz and the estimated volume of raw data is about 2 PB/year. But the event rate of reactor anti-neutrino is only about 60/day. So one of challenges...

53. Quantum control algorithm of imperfect knowledge bases of intelligent cognitive controllers

Проф. Sergey Ulyanov (professor)

06.07.2021, 13:30

8. Quantum information processing

Sectional reports

Quantum information processing

The quantum self-organization algorithm model of wise knowledge base design for intelligent fuzzy controllers with required robust level considered. Background of the model is a new model of quantum inference based on quantum genetic algorithm. Quantum genetic algorithm applied on line for the quantum correlation’s type searching between unknown solutions in quantum superposition of imperfect...

116. Visualization of Experimental Data in Web-based Virtual Reality

Štefan Korečko (DCI FEEI TU Košice, Slovakia), Martin Vala (JINR), Mr Martin Fekete (Department of Computers and Informatics, Faculty of Electrical Engineering and Informatics, Technical University of Košice)

06.07.2021, 13:30

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

Technological advances in the field of virtual reality and personal computation in general brought us to the era of web-based virtual reality, where virtual environments can be accessed directly from web browsers and without the need of installation of any additional software. Such online virtual environments seem to be a promising tool for scientific data visualization. When accessed through...

161. Design and development of application software for MPD distributed computing infrastructure

Igor Pelevanyuk (Joint Institute for Nuclear Research)

06.07.2021, 13:45

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

Multi-Purpose Detector collaboration began using distributed computing for centralized Monte-Carlo generation in the mid of 2019. DIRAC Interware is used as a platform for the integration of heterogeneous distributed computing resources. Since that time workflows of job submission, data transfer, and storage were designed, tested, and successfully applied. Moreover, we observe the growth of...

108. Direct computational experiment in storm hydrodynamics of marine objects

Alexander Degtyarev (Professor), Ivan Gankevich (Saint Petersburg State University)

06.07.2021, 13:45

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

The paper presents and discusses a new computer toolkit for assessing the seaworthiness of a ship in stormy sailing conditions, intended for testing new design solutions for promising ocean-going ships and ships of unlimited ocean navigation, as well as organizing full-function simulators. The presented toolkit can be used by captains to select effective and safe sailing modes, as well as to...

68. Mechanisms for identifying the patterns of the dynamics of scientific and technical publications on the example of the thematic direction "Robotics"

Mr Andrey Cherkasskiy (National Research Nuclear University MEPhI (Moscow Engineering Physics Institute))

06.07.2021, 13:45

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Introduction
In world practice, the number of published articles in leading scientific publications is indicators of the results of scientific activities of researchers, research organizations and higher educational institutions. International publication activity reflects the level of development of national science against the background of other countries, especially in the field of basic...

92. Quantum Machine Learning for HEP detectors simulations

Florian Rehm (CERN, RWTH Aachen University (DE))

06.07.2021, 13:45

8. Quantum information processing

Sectional reports

Quantum information processing

Quantum Machine Learning is one of the most promising applications on near-term quantum devices which possess the potential to solve problems faster than traditional computers. Classical Machine Learning is taking up a significant role in particle physics to speed up detector simulations. Generative Adversarial Networks (GANs) have scientifically proven to achieve a similar level of accuracy...

186. Data analysis platform for stream and batch data processing on hybrid computing resources

Ivan Kadochnikov (JINR, PRUE)

06.07.2021, 14:00

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The modern Big Data ecosystem provides tools to build a flexible platform for processing data streams and batch datasets. Supporting both the functioning of modern giant particle physics experiments and the services necessary for the work of many individual physics researchers generate and transfer large quantities of semi-structured data. Thus, it is promising to apply cutting-edge...

109. Development of a tool for interactive detailing of areas of objects for the strength modeling system

Egor Budlov

06.07.2021, 14:00

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Introduction

The development of technological progress makes increased demands on the strength properties of structural elements of buildings and structures, machines and mechanisms, and a decrease in their material consumption.

This leads to the need for effective use of existing and creation of new methods of solid mechanics and the training of new highly qualified specialists.
One of...

86. EFFECTIVE ALGORITHM OF CALCULATING THE WIGNER FUNCTION FOR A QUANTUM SYSTEM WITH A POLYNOMIAL POTENTIAL

Evgeny Perepelkin (JINR)

06.07.2021, 14:00

8. Quantum information processing

Sectional reports

Quantum information processing

When considering quantum systems in phase space, the Wigner function is used as a function of quasidensity of probabilities. Finding the Wigner function is related to the calculation of the Fourier transform from a certain composition of wave functions of the corresponding quantum system. As a rule, knowledge of the Wigner function is not the ultimate goal, and calculations of mean values...

159. Web based Event Display server for MPD/NICA experiment

Alexander Krylov

06.07.2021, 14:00

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

There are different methods for monitoring engineering, network, and computer systems for high-energy physics experiments. As a rule, they have a common name - Event Display and include a whole range of monitoring and control systems. During the experiment, the facility operator should receive comprehensive information about detector performance in a well understandable and intuitive form to...

45. Classical Fisher information for the state space of N-level systems through the Wigner function

Vahagn Abgaryan (JINR LTP)

06.07.2021, 14:15

8. Quantum information processing

Sectional reports

Quantum information processing

The studies of the geometrical aspects of the quantum information grow very actual owing to practical purposes. Due to a request coming from the quantum technology, formulation of the quantum estimation theory turn to be in the frontier of a modern research. Particularly, the issue of interrelations between the phase space quasidistributions and classical Fisher metric are of current...

142. Detection of fertile soils based on satellite imagery processing

Dr Valery Grishkin (Saint Petersburg State University)

06.07.2021, 14:15

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

The paper proposes a method for detecting fertile soils based on the processing of satellite images. As a result of its application, a map of the location of fertile and infertile soils for a given region of the earth's surface is formed and the corresponding areas are calculated. Currently, data from most satellites are in the public domain and, as a rule, are multispectral images of the...

113. Intelligent Networks: using convolutional LSTM models to estimate network traffic

Ms Joanna Waczynska (Wroclaw University of Science and Technology)

06.07.2021, 14:15

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The Large Hadron Collider experiments at CERN produce a large amount of data which are analyzed by the High Energy Physics (HEP) community in hundreds of institutes around the world.
Both efficient transport and distribution of data across HEP centres are, therefore, crucial.
The HEP community has thus established high-performance interconnects for data transport---notably the Large Hadron...

77. Performance Analysis and Optimization of MPDroot

Slavomir Hnatic (JINR)

06.07.2021, 14:15

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

We present analysis of performance of MPD data analysis/simulation software MPDroot by profilers and benchmarks.
Based on this we draw preliminary conclusions and set perspectives for future optimization improvements.

91. Benchmark of Generative Adversarial Networks for Fast HEP Calorimeter Simulations

Florian Rehm (CERN, RWTH Aachen University (DE))

06.07.2021, 14:30

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Accurate simulations of elementary particles in High Energy Physics (HEP) detectors are fundamental to accurately reproduce and interpret the experimental results and to correctly reconstruct particle flows. Today, detector simulations typically rely on Monte Carlo-based methods which are extremely demanding in terms of computing resources. The need for simulated data at future experiments -...

44. Describing quantumness of qubits and qutrits by Wigner function’s negativity

Astghik Torosyan (LIT)

06.07.2021, 14:30

8. Quantum information processing

Sectional reports

Quantum information processing

According to modern views, the Wigner quasiprobability distribution provides a qualitative information on many quantum phenomena occurring in diverse physical systems. The Wigner function has all the properties of statistical distributions except one: taking negative values for some quantum states, the Wigner function turns to be not a proper distribution, and hence it indicates the existence...

131. Potential of Neural Networks for Air Quality Sensor Data Processing and Analysis

Jan Bitta (VSB-TU Ostrava)

06.07.2021, 14:30

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Air quality sensors represent an emerging technology for air monitoring quality. Their main advantage is that they are significantly cheaper monitoring devices compared to standard monitoring equipment. Low-cost, mass-produced sensors have a potential to form much denser monitoring networks and provide more detailed information on air pollution distribution. The drawback of sensor air...

152. Многопоточный режим моделирования событий в пакете BmnRoot

Stepanova Margarita (SPbSU)

06.07.2021, 14:30

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

Для исследований на ускорительном комплексе NICA (ОИЯИ) необходимы эффективные и быстрые программные реализации алгоритмов моделирования и реконструкции событий. Созданный для эксперимента BM@N пакет BmnRoot базируется на среде ROOT, GEANT4 и объектно-ориентированном фреймворке FairRoot и включает инструменты для исследования характеристик детектора BM@N, а также восстановления и анализа...

132. Air Pollution Modelling Using Spatial Analysis and Neural Networks

Dr Vladislav Svozilík (JINR LIT)

06.07.2021, 14:45

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

In a huge number of applications, air pollution dispersion modelling using standard Gaussian methodologies is an excessively data-intensive process that requires considerable computing power. Land Use Regression (LUR) represents an alternative modelling methodology. LUR presumes that pollution concentration is determined by factors obtained via spatial analysis. These factors are chosen on the...

101. Identification of news text corpora influencing the volatility of financial instruments

alexey Stankus (-)

06.07.2021, 14:45

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Identifying news that affects financial markets is an important task on the way to predicting financial markets. A large number of articles are devoted to this topic. But the main problem for analyzing news is neural networks what used. These neural networks are created to analyze user reports about a particular object, be it a restaurant, a movie or a purchased item. In such reports, the...

184. On the geometry of non-maximal strata qudit space with Bures metric

Martin Bures (IEAP, CTU Prague, Czechia & JINR Dubna, Russia)

06.07.2021, 14:45

8. Quantum information processing

Sectional reports

Quantum information processing

Modern applications of quantum mechanics renewed interest in the properties of the set of density matrices of finite size. The issue of establishing of Riemannian structures on the quantum counterparts of space of probability measures became a subject of recent investigations.
We study quantum analogues of a well-known, natural Riemannian metric, the so-called Fisher metric. Explicit...

59. Classification of lung X-rays with pneumonia disease using deep learning models

Проф. EUGENE SHCHETININ (Financial University under the Government of Russian the Federation)

06.07.2021, 15:20

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Pneumonia is a life-threatening lung disease caused by either a bacterial or viral infection. It can be life-threatening if not acted on at the right time, and so early diagnosis of pneumonia is vital. The aim of this work is the automatic detection of bacterial and viral pneumonia on the basis of X-ray images. Four different pre-trained deep convolutional neural networks (CNN): VGG16,...

156. Efficient gossip-based protocol in the Neo blockchain network

Anna Shaleva

06.07.2021, 15:20

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Epidemic algorithms are widely explored in the case of distributed systems based on trustful environments. However, an assumption on arbitrary peers behaviour in Byzantine fault tolerance problem calls into question the appropriateness of well-studied gossip algorithms since some of them are based on aggregated network information, i.e. the number of nodes in the network, etc. Given this...

83. Missing Mass Method for reconstruction of short-lived particles in the CBM and STAR experiments

Pavel Kisel (Frankfurt Uni, JINR)

06.07.2021, 15:20

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The search for short-lived particles is an important part of the physics research in experiments with relativistic heavy ions.
Such investigations mainly study decays of neutral particles into charged daughter particles, which can be already registered in the detector system. In order to find, select and study the properties of such short-lived particles in real time in the CBM experiment...

126. National Research and Education Network of Russia: directions of development in the context of expanding of international cooperation

Dr Alexey Abramov (Joint SuperComputer Center of the Russian Academy of Science)

06.07.2021, 15:20

2. Research infrastructure

Sectional reports

The report overviews the current state and key directions of the advanced development of the National Research and Education Network (NREN) of Russia for the period of 2021-2024. Unified NREN called National

Research Computer Network (NIKS) was created in 2019 according to the results of integration of federal-level telecommunication networks in the fields of higher education (RUNNet) and...

154. Comparative analysis and applicability determination for several DLT solutions

Mr Anar Faradzhov (Saint Petersburg State University)

06.07.2021, 15:35

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Potential benefits of implementation of distributed ledger technology are widely discussed among different business actors and governmental structures. Within the last decade, with growing popularity of blockchain-based payment systems and cryptocurrencies, these discussions considerably sharpened. Therefore, an extensive body of research has emerged on this soil. The goal of this study is to...

148. Forecasting and assessment of land conditions using neural networks

Ms Anastasiia Khokhriakova, Mr Valery Grishkin

06.07.2021, 15:35

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

This paper proposes a method for predicting and assessing land conditions based on satellite image processing using neural networks. In some regions, mainly based on agriculture and cattle breeding, the threat of irreversible soil changes has appeared, in particular desertification, which can lead to serious environmental and economic problems. Therefore, it is necessary to identify both the...

157. Neural Networks in Modeling Beam Dynamics using Taylor Mapping

Nataliia Kulabukhova (Saint Petersburg State University)

06.07.2021, 15:35

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The paper describes method for modeling beam dynamics based on the calculation of ordinary differential equations with Taylor mapping. This method allows you to get the solutions of the system both in symbolic and numerical form. Using numerical simulation methods, one can obtain partial solutions of beam dynamics process. The paper considers the possibility of solving the inverse problem -...

119. The concept of training IT professionals in cross-cutting digital technologies

Проф. Evgenia Cheremisina (Dubna State University)

06.07.2021, 15:35

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

The formation of a new generation of digital technologies, which were called «cross-cutting» due to the scale and depth of their influence, determined a large-scale transformation of business and social models. These changes have a strong impact on the content of professional activity: new skills are required from employees, and therefore new competencies. The rapid digitalization of the...

170. Deep learning for automatic RF-modulation classification

M. Dima (University of Bucharest)

06.07.2021, 15:50

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Classical methods use statistical-moments to determine the type of modulation in question. This essentially correct approach for discerning amplitude modulation (AM) from frequency modulation (FM), fails for more demanding cases such as AM vs. AM-LSB (lower side-band rejection) - radio signals being richer in information that statistical moments. Parameters with good...

176. Grammar parser-based solution for the description of the computational graph within GNA framework

Nikita Tsegelnik (BLTP, JINR)

06.07.2021, 15:50

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

The data flow paradigm has established itself as a powerful approach to the machine learning. Indeed, it is also very powerful for the computational physics, although it is not used as much in the field. One of the complications is that physical models are much less homogeneous compared to ML, which makes their description a quite complicated task.
In this talk we present a syntax analyzer...

162. Using distributed computing systems to solve the problem of image classification using deep neural networks

Ilya Kurochkin (IITP RAS)

06.07.2021, 15:50

4. Distributed computing applications

Sectional reports

Big data Analytics and Machine learning.

Machine learning methods and, in particular, deep neural networks are often used to solve the problem of image classification. There is a tendency to increase the amount of training data and the size of neural networks. The process of training a deep neural network with millions parameters can take hundreds of hours on modern computing nodes. Parallel and distributed computing can be used to...

144. Обучение без учителя на данных узкой направленности агрегированных автоматически.

Ekaterina Pavlova (Saint Petersburg State University)

06.07.2021, 15:50

9. Big data Analytics and Machine learning

Sectional reports

Информация, публикуемая пользователями в открытом доступе, может служить хорошим ресурсом для сбора данных при формировании датасетов для обучения нейронных сетей. Одной из самых больших существующих платформ для обмена фотографиями и видеозаписями является Instagram. Основным методом взаимодействия пользователей друг с другом на данной платформе является публикация изображений. При этом...

145. High resolution image processing and land cover classification for hydro-geomorphological high-risk area monitoring

Giorgia Miniello (UNIVERSITA' DEGLI STUDI DI BARI E INFN BARI)

06.07.2021, 16:05

4. Distributed computing applications

Sectional reports

Big data Analytics and Machine learning.

High-resolution images processing for land-surface monitoring is fundamental to analyse the impact of different geomorphological processes on earth surface for different climate change scenarios. In this context, photogrammetry is one of the most reliable techniques to generate high-resolution topographic data, being key to territorial mapping and change detection analysis of landforms in...

160. Improvements of the LOOT model for primary vertex finding based on the analysis of development results

Ekaterina Rezvaya

06.07.2021, 16:05

3. Computing for MegaScience Projects

Computing for MegaScience Projects

The recognition of particle trajectories (tracks) from experimental measurements plays a key role in the reconstruction of events in experimental high-energy physics. Knowledge about the primary vertex of an event can significantly improve the quality of track reconstruction. To solve the problem of primary vertex finding in the BESIII inner tracking detector we applied the LOOT program which...

169. NARX neuromorphic software in ECG wave prediction

T. Dima (University of Bucharest)

06.07.2021, 16:05

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

We present an approach to predict ECG waves with non-linear autoregressive exogenous neuromorphic (NARX) software. These predictions are important in comparing the underlying QRS complex of the ECG-wave with the slowly deteriorating waves (or arrythmia) in cardiac patients. A deep Q-wave for instance (such as 1/4 of the R-wave) is a typical sign of (inferior wall) myocardial...

171. Using data from the labor market for analysis and education

Irina Filozova (JINR MLIT, Dubna State University, Plekhanov Russian University of Economics)

06.07.2021, 16:05

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Education systems provide specialists of different levels and specialization for the labor market. However, in the modern dynamic world of artificial intelligence, pandemic, and remote work, the labor market evolves dramatically from year to year. Universities and colleges must keep track of these changes to adapt educational programs and manage the number of student slots offered for...

85. Explanation of NMR mobility of peptide dendrimers using distributed computing

Oleg Shavykin (St. Petersburg State University, 7/9 Universitetskaya nab., 199034 St. Petersburg, Russia; ITMO University, Kronverkskiy pr. 49, 197101 St. Petersburg, Russia; Tver State University, Zhelyabova 33, Tver, Russia)

06.07.2021, 16:20

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

We applied distributed computing to study new peptide dendrimers with Lys-2Lys and Lys-2Arg repeating units in water. These molecules are promising nanocontainers for the drug and gene delivery. The dendrimers have recently been synthesized and studied by NMR (Sci. Reports, 2018, 8, 8916; RSC Advances, 2019, 9, 18018) and successfully tested as carriers for gene delivery (Bioorg. Chem., 2020,...

115. SQL query execution optimization on Spark SQL

Gleb Mozhaiskii

06.07.2021, 16:20

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The Spark – Hadoop ecosystem includes a wide variety of different components and can be integrated with any tool required for Big Data nowadays. From release-to-release developers of these frameworks optimize the inner work of components and make their usage more flexible and elaborate.
Anyway, since inventing MapReduce as a programming model and the first Hadoop releases data skew was and...

42. The analysis of the educational measurement results, and its providing as “software-as-a-service” solution in eLearning

Ms Julia Lavdina (National Research Nuclear University MEPhI (Moscow Engineering Physics Institute))

06.07.2021, 16:20

4. Distributed computing applications

Sectional reports

In modern eLearning systems, educational measurements are used both to evaluate the students’ achievements, and to control the learning process. However, eLearning systems usually have comparatively trivial embedded features for analyzing measurement results, which are not of considerable interest for sufficient statistical research of the assessment tools quality. To identify the...

165. TrackNETv3 with optimized inference for BM@N tracking

Anastasiia Nikolskaia

06.07.2021, 16:20

3. Computing for MegaScience Projects

Sectional reports

Computing for MegaScience Projects

There are local and global approaches to do track reconstruction depending on the amount of input data available for training a neural network model that solves the reconstruction problem. Global methods need access to all tracks in an event that results in a high memory footprint. We have successfully applied the recurrent neural network (RNN) TrackNETv2 and its updated version v2.1 to the...

27. Fractal thermodynamics, big data and its 3D visualization

Victor Tsvetkov (Tver State University)

06.07.2021, 16:35

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

The need for big data analysis takes place in many areas of science and technology: economics, medicine, geophysics, astronomy, particle physics and many others.
This task is greatly simplified if big data has structural patterns. In this talk, we will consider the case when big data with a high degree of accuracy are fractals.
We propose to analyze the fractal structure of big data based on...

220. Joint Scientific and Educational projects of JINR and NOSU

Nelli Pukhaeva (JINR / NOSU)

06.07.2021, 16:35

Sectional reports

153. Extraction of traffic features in Software Defined Networks using an SDN Controller

Mr Sergey Volkov (Peoples' Friendship University of Russia (RUDN University); Federal Research Center "Computer Science and Control" RAS)

06.07.2021, 16:50

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Machine learning methods can be used to solve the problems of detecting and countering attacks on software-defined networks. For such methods, it is necessary to prepare a large amount of initial data for training. Mininet is used as a modeling environment for SDN. The main tasks of modeling a software-defined network are studying traffic within the network, as well as practicing various...

93. Разработка и исследование распределенных алгоритмов управления системами роевого интеллекта

Mr Артем Горемыкин (университет "Дубна")

06.07.2021, 17:05

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Настоящая работа посвящена разработке и исследованию методов управления коллективным поведением в роевых робототехнических системах на примере решения модельной задачи уборки роем роботов заданной ограниченной территории. В работе рассматривается несколько распределенных алгоритмов решения поставленной задачи, основанных на различных классических методах и моделях роевого интеллекта....

139. Применение методов машинного обучения для задачи распознавания русских дореволюционных печатных текстов

Mr Владислав Федоров (ф-т ВМК МГУ им. М.В. Ломоносова)

06.07.2021, 17:20

9. Big data Analytics and Machine learning

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Настоящая работа посвящена вопросам применения технологий оптического распознавания символов и методов машинного обучения для распознавания печатных русскоязычных текстов XIX века. Анализируются особенности данной задачи по сравнению с общей задачей оптического распознавания символов. Проводится обзор существующих методов и программ для решения рассматриваемой проблемы. Предлагается свой...

210. Optimization of the computation of the multidimentional integrals for estimation of the meson lifetime

Daviti Goderidze

06.07.2021, 17:35

10. Distributed computing, HPC and ML for solving applied tasks

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

To calculate the lifetime of mesons in hot and dense nuclear matter, it is necessary to computate the 5-dimentional integrals with complicated integrand function. This work presents algorithms and methods for calculating complicated integrals based on the Monte-Carlo method. For optimization of computation the algorithm of parallel calculations was implemented in C++ programming language...

194. From Quantum Speed-up to Supremacy and Advantage

Cristian Calude (University of Auckland)

07.07.2021, 09:00

Plenary reports

Plenary reports

Quantum computing began in the early 1980s when physicist Paul Benioff constructed a quantum mechanical model of Turing machine and physicist Richard Feynman and mathematician YuriManin discussed the potential of quantum computers to simulate phenomena a classical computer could not feasibly do.
In 1994 Peter Shor developed a polynomial quantum algorithm for factoring integers with the...

204. High-performance quantum computing technologies

Alexey Fedorov (Russian Quantum Center)

07.07.2021, 09:45

Plenary reports

Plenary reports

181. Clustering in ontology-based exploratory analysis of scientific productivity

Pawel Lula (Cracow University of Economics, Poland)

07.07.2021, 10:50

9. Big data Analytics and Machine learning

Sectional reports

Plenary reports

Ontology-based approach in exploratory analysis of textual data can significantly improve the quality of the obtained results. On the other hand, the use of domain knowledge defined in the form of ontologies increases the time needed to prepare a model and makes required calculations more complex. The presentation will discuss selected aspects of cluster analysis performed on documents...

168. Серверы Dell EMC PowerEdge ЛОКОМОТИВ ИТ-ИННОВАЦИЙ

Nikita Stepanov (Dell)

07.07.2021, 11:30

2. Research infrastructure

Plenary reports

Айтикост и Dell в понедельник . доклад спонсоров

208. Опыт и возможности Softline в обеспечении инфраструктуры для научных исследований

Mr Сергей Монин (Софтлайн)

07.07.2021, 12:00

Plenary reports

Plenary reports

221. Вычислительные технологии класса энтерпрайз

Valery Yegorshev (NIAGARA COMPUTERS, LLC)

07.07.2021, 12:30

Plenary reports

197. RDIG-M

Vasiliy Velikhov (Kurchatov Institute National Research Centre)

07.07.2021, 13:20

1. Distributed computing systems

Plenary reports

Plenary reports

-

189. Current status of the MICC: an overview

Tatiana Strizh (JINR)

08.07.2021, 09:00

Plenary reports

Plenary reports

-

78. PIK Data Centre status

Andrey Kiryanov (PNPI)

08.07.2021, 09:45

5. High Performance Computing

Plenary reports

Plenary reports

In the framework of the PIK nuclear reactor reconstruction project, a PIK Data Centre was commissioned in 2017. While the main purpose of the Centre is storage and processing of PIK experiments data, its capacity is also used by other scientific groups at PNPI and outside for solving problems in different areas of science. PIK Data Centre is an integral part of computing facilities of NRC...

57. Offline Software and Computing for the SPD experiment

Alexey Zhemchugov (JINR)

08.07.2021, 11:00

3. Computing for MegaScience Projects

Plenary reports

Plenary reports

The SPD (Spin Physics Detector) is a planned spin physics experiment in the second interaction point of the NICA collider that is under construction at JINR. The main goal of the experiment is the test of basic of the QCD via the study of the polarized structure of the nucleon and spin-related phenomena in the collision of longitudinally and transversely polarized protons and deuterons...

188. IT solutions for JINR tasks on the “Govorun” supercomputer

Maxim Zuev (JINR)

08.07.2021, 11:45

Plenary reports

Plenary reports

The “Govorun” supercomputer is a heterogeneous computing system that contains computing architectures of different types, including graphics accelerators. The given architecture of the supercomputer allows users to choose optimal computing facilities for solving their tasks.
To enhance the efficiency of solving user tasks, as well as to expand the efficiency of utilizing both the computing...

183. Population annealing method and hybrid supercomputer architecture

Lev Shchur (leading reseacher, Landau Institute for Theoretical Physics)

08.07.2021, 13:30

5. High Performance Computing

Sectional reports

HPC

A population annealing method is a promising approach for large-scale simulations because it is potentially scalable on any parallel architecture. We report an implementation of the algorithm on a hybrid program architecture combining CUDA and MPI [1]. The problem is to keep all general-purpose graphics processing unit devices as busy as possible by efficiently redistributing replicas. We...

82. Russian data lake prototype as an approach towards national federated storage for Megascience.

Andrey Kiryanov (PNPI)

08.07.2021, 13:30

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

A substantial data volume growth will appear with the start of the HL-LHC era. It is not well covered by the current LHC computing model, even taking into account the hardware evolution. The WLCG DOMA project was established to provide data management and storage researches. National data lake r&d's, as a part of the DOMA project, should address the study of possible technology solutions for...

138. Some Aspects of the Workflow Scheduling in the Computing Continuum Systems

Mr Vladislav Kashansky (University of Klagenfurt and South Ural State University)

08.07.2021, 13:30

1. Distributed computing systems

Distributed computing systems

Contemporary computing systems are commonly characterized in terms of data-intensive workflows, that are managed by utilizing large number of heterogeneous computing and storage elements interconnected through complex communication topologies. As the scale of the system grows and workloads become more heterogeneous in both inner structure and the arrival patterns, scheduling problem becomes...

62. Система массовой памяти МИВК ОИЯИ. Статус, перспективы.

Vladimir Trofimov (JINR)

08.07.2021, 13:30

2. Research infrastructure

Research infrastructure

Развитие экспериментов в различных областях приводит к увеличению объёма хранения и интенсивности обработки данных. Этот факт приводит к повышению количественных и качественных требований к системам хранения данных. Рассматриваются текущий статус, результаты проведённой в последнее время модернизации и перспективы дальнейшего развития работизированного хранилища данных Многофункционального...

66. LOAD BALANCING STRATEGIES IN GRID-SYSTEMS

Mr Sergei A. Mamaev (Data center engineer, RU-CENTER company)

08.07.2021, 13:45

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The research of load balancing strategies in Grid systems is carried out. The main classes of load distribution strategies are identified with the aim of possibly increasing the efficiency of distributed systems. A model based on the fractal method for describing the dynamics of the load is considered.

164. Multi-GPU training and parallel CPU computing for the machine learning experiments using Ariadne library

Egor Shchavelev (Saint Petersburg State University)

08.07.2021, 13:45

5. High Performance Computing

HPC

Modern machine learning (ML) tasks and neural network (NN) architectures require huge amounts of GPU computational facilities and demand high CPU parallelization for data preprocessing. At the same time, the Ariadne library, which aims to solve complex high-energy physics tracking tasks with the help of deep neural networks, lacks multi-GPU training and efficient parallel data preprocessing on...

150. The JINR distributed information and computing environment: participants, features and challenges

Nikolay Kutovskiy (JINR)

08.07.2021, 13:45

1. Distributed computing systems

Sectional reports

Distributed computing systems

The JINR distributed information and computing environment (DICE) was created to join resources for solving common scientific tasks as well as to distribute a peak loads across resources of partner organizations from JINR Member States. To monitor hardware resources and services of growing DICE infrastructure a system based on Prometheus and Thanos was designed and deployed. Collected metrics...

87. Usage of time series databases in the Grafana platform for the NETIS service

Evgeny Alexandrov (JINR)

08.07.2021, 13:45

2. Research infrastructure

Sectional reports

Research infrastructure

NetIs is a service used to monitor the Data Acquisition network of the ATLAS experiment. The first version was developed at CERN in 2010. The network group would like to replace NetIs because it is difficult to maintain and store data in a Round Robin Database (RRD), resulting in a loss of granularity over time that makes the tool unsuitable for retrieving accurate values from the past. The...

61. APPROACH TO REMOTE PARTICIPATION IN THE ITER EXPERIMENTAL PROGRAM. EXPERIENCE FROM MODEL OF RUSSIAN REMOTE PARTICIPATION CENTER

Mr Oleg Semenov (Project Center ITER)

08.07.2021, 14:00

6. Data Management, Organisation and Access

Plenary reports

Data Management, Organization and Access

The model of Russian Remote Participation Center (RPC) was created under the contract between Russian Federation Domestic Agency (RF DA) and ROSATOM as the prototype of full-scale Remote Participation Center for ITER experiments and for coordination activities in the field of Russian thermonuclear research. This prototype was used for investigation of the following technical and scientific...

173. Implementing the Graph Model of the Spread of a Pandemic on GPUs

Vladimir Sudakov (Plekhanov Russian University of Economics, Keldysh Institute of Applied Mathematics (Russian Academy of Sciences))

08.07.2021, 14:00

5. High Performance Computing

Sectional reports

HPC

Modeling the spread of viruses is an urgent task in modern conditions. In the created model, contacts between people are represented in the form of the Watz and Strogatz graph. We studied graphs with tens of thousands of vertices with a simulation period of six months. The paper proposes methods for accelerating computations on graph models using graphics processors. In the considered problem,...

50. Monitoring System for the Russian Scientific Data Lake Prototype

Mr Aleksandr Alekseev (Ivannikov Institute for System Programming of the RAS)

08.07.2021, 14:00

2. Research infrastructure

Sectional reports

Research infrastructure

The Russian Scientific Data Lake is a part of Data Lake R&D conducted by the DOMA project. It aims to mitigate the present LHC computing model drawbacks to cope with an unprecedented scientific data volume at the multi-exabyte scale that will be delivered by experiments in the High Luminosity phase of the LHC. The prototype of the Russian Scientific Data Lake is being implemented and it tests...

129. Usage of JINR SSO authentication and authorization system with distributed data processing services.

Andrey Iachmenev

08.07.2021, 14:00

1. Distributed computing systems

Sectional reports

Distributed computing systems

The amount of data produced by a scientific community already measured tens and hundreds of petadytes and will significantly grow in future. Distributed computing systems have proven to be an effective solution for handling such data streams. Technologies and a set of products that allow deployment and use of important components of a distributed system already existed, quite stable and...

67. Developing a Toolkit for Task Characteristics Prediction Based on Analysis of Queue’s History of a Supercomputer

Mr Mahdi Rezaei (Moscow Institute of Physics and Technology, Moscow, Russia)

08.07.2021, 14:15

5. High Performance Computing

HPC

Empirical studies have repeatedly shown that in High-Performance Computing (HPC) systems, the user's resource estimation lacks accuracy [1]. Therefore, resource underestimation may remove the job at any step of computing, and subsequently allocated resources will be wasted. Moreover, resource overestimation also will waste resources. The SLURM, a famous job scheduler, has a mechanism to...

47. JINR CMS Tier-1,LCG2 accounting system

Ivan Kashunin (JINR)

08.07.2021, 14:15

2. Research infrastructure

Sectional reports

Research infrastructure

The problem of evaluating the efficiency of the JINR MLIT grid sites has always been topical. At the beginning of 2021, a new accounting system was created, it managed to fully cover the functionality of the previous system and further expand it. The report will provide detailed information on the implemented accounting system.

Проблема оценки эффективности работы грид сайтов ЛИТ...

141. Modeling the process of executing an evolutionary algorithm on the desktopgrid

Nikolay Khrapov (Pavlovich)

08.07.2021, 14:15

1. Distributed computing systems

Sectional reports

Distributed computing systems

The topic of the presented report is various approaches to modeling the process of solving optimization problems using the desktopgrid [1]. The report summarizes the practical experience of performing computations on local infrastructures and on voluntary computing projects. The creation of preliminary models of the computational process will allow to avoid many systemic complexities in the...

64. THE ALGORITHM FOR SOLVING THE PROBLEM OF SYNTHESIS OF THE OPTIMAL LOGICAL STRUCTURE OF DISTRIBUTED DATA IN ARCHITECTURE OF GRID SERVICE

Elena Nurmatova (Russia)

08.07.2021, 14:15

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

Abstract. The questions of constructing optimal logical structure of a distributed database (DDB) are considered. Solving these issues will make it possible to increase the speed of processing requests in DDB in comparison with a traditional database. In particular, such tasks arise for the organization of systems for processing huge amounts of information from the Large Hadron Collider – the...

38. DEVELOPMENT OF EFFECTIVE ACCESS TO THE DISTRIBUTED SCIENTIFIC AND EDUCATIONAL E-INFRASTRUCTURE

Grigore Secrieru (Vasile)

08.07.2021, 14:30

2. Research infrastructure

Sectional reports

Research infrastructure

The article describes approaches to the modernization of a distributed electronic infrastructure that combines various types of resources aimed at supporting the research **and educat**ional activities in Moldova. The development trends of computer infrastructures and technologies aimed at creating conditions for solving complex scientific problems with high requirements for computing...

79. IT for air quality management - mathematical modeling verified by special sampling and nuclear analytical methods and Air Pollution Management System (AQMS)

Petr Jancik (JINR; VSB - Technical University of Ostrava)

08.07.2021, 14:30

4. Distributed computing applications

Sectional reports

Distributed computing systems

AQMS - in this case, the broad name includes in particular measurement and mathematical modeling of air pollution, geoinformation technologies for the analysis of their results and preparation and implementation of

modeling on parallel supercomputer clusters. In the past few years, my team and I have been researching and refining mathematical models, expanding the amount of processed input...

130. Research of improving the performance of explicit numerical methods on the x86 and ARM CPU

Vladislav Furgailo

08.07.2021, 14:30

5. High Performance Computing

Sectional reports

HPC

Explicit numerical methods are used to solve and simulate a wide range of mathematical problems whose origins can be mathematical models of physical conditions. However, simulations with large model spaces can require a tremendous amount of floating point calculations and run times of several months or more are possible even on large HPC systems.
The vast majority of HPC systems in the field...

167. The development of a new conditions database prototype for ATLAS RUN3 within the CREST project

Mikhail Mineev (JINR)

08.07.2021, 14:30

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The CREST project for a new conditions database prototype for Run3 (intended to be used for production in Run 4) is focused on improvement of Athena based access, metadata management and, in particular, global tag management. The project addresses evolution of the data storage design and conditions data access optimization, enhancing the caching capabilities of the system in the context of...

124. Analysis of the effectiveness of various methods for parallelizing data processing implemented in the ROOT package.

Tatyana Solovjeva (Jinr)

08.07.2021, 14:45

5. High Performance Computing

Sectional reports

HPC

The ROOT software package has a central role in high energy analytics and is being upgraded in several ways to improve processing performance. In this paper, we will consider several tools implemented in this framework for calculations on modern heterogeneous computing architectures.

PROOF (Parallel ROOT Facility – an extension of ROOT system) uses the natural parallelism of data structures...

51. Data Center Simulation for the BM@N experiment of the NICA Project

Daria Priakhina (ЛИТ)

08.07.2021, 14:45

2. Research infrastructure

Research infrastructure

One of the uppermost tasks in creating a computing system of the NICA complex is to model centers of storing and processing data that come from experimental setups of the complex, in particular, the BM@N detector, or are generated using special software for checking of the developed data processing algorithms and for comparison with the expected physical result.

After reviewing the existing...

134. Development of the Condition Database for the experiments of the NICA complex

Konstantin Gertsenberger (JINR)

08.07.2021, 14:45

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

Processing and analyzing of experimental and simulated data are an integral part of all modern high-energy physics experiments. These tasks are of particular importance in the experiments of the NICA project at the Joint Institute for Nuclear Research (JINR) due to the high interaction rate and particle multiplicity of ion collision events, therefore the task of automating the considered...

111. The graph diameter of a distributed system with a given dominant set

Ilya Kurochkin (IITP RAS)

08.07.2021, 14:45

1. Distributed computing systems

Sectional reports

Distributed computing systems

In this work consider a distributed computing system in which the control functions are dispersed in several dominant nodes that are directly connected to all the others. This configuration reduces the vulnerability of the entire network, since the failure of a single control element immediately disrupts its operation. On the other hand, the large length of the maximum shortest chain...

104. A virtual testbed for optimizing the performance of a new type of accelerators

Maria Mingazova (St.Petersburg State University)

08.07.2021, 15:30

2. Research infrastructure

Sectional reports

Research infrastructure

The concept of a "virtual testbed", namely the creation of problem-oriented environment for full-featured modeling for the investigated phenomenon or behavior of a complex technical object,
has now acquired a finished look. The design of this concept contributed to the development of complex mathematical models suitable for full-fledged computational experiments and improvement of a computer...

81. Development of the Event Metadata System for the NICA experiments

Peter Klimai (INR RAS)

08.07.2021, 15:30

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

Particle collision experiments are known to generate substantial amount of data that must be stored and, later, analyzed. Typically, only a small subset of all the collected events is relevant when performing a particular physics analysis task. Although it is possible to obtain the required subset of records directly, by iterating through the whole volume of the collected data, the process is...

222. Intel oneAPI for xPU

Dmitry Sivkov (Intel)

08.07.2021, 15:30

HPC

40. Jira plugin for ALICE instance

Andrey Kondratyev (JINR)

08.07.2021, 15:30

1. Distributed computing systems

Sectional reports

Distributed computing systems

As a bug tracking and project management system, the ALICE experiment uses Jira software, which provides a wide range of configuration options. Jira also allows to significantly expand the basic functionality with custom plugins. In this work, the LinkedIssuesHasStatus plugin for the JIRA service of the ALICE experiment is developed and implemented. Plugin returns tickets that have linked...

80. New hardware testing methodology at IHEP data center

Victoria Ezhova (IHEP)

08.07.2021, 15:45

2. Research infrastructure

Sectional reports

Research infrastructure

The modern computing center is not only a production capacity, it is also a stable work. Stability means not only software, the reliability of the software component. It is also assembly components. It is important to ensure this before the production. One of the ways is testing components for performance, reliability, assembly defects. In this work we will present the methodology for...

105. Remote Procedure Call protocol with support for higher-order functions

Fedor Bukreev (mipt-npm)

08.07.2021, 15:45

1. Distributed computing systems

Distributed computing systems

We are developing a tool for for calling functions between different *environments*. And by "different environment" we mean both, different programming languages and different machines. This tool is a remote procedure call protocol(and it's implementation), that is optimized for simplicity and can support higher-order functions. In our implementation, functions are never serialized and are...

122. Tape libraries as a part of JINR MICC mass storage system.

Алексей Голунов (ЛИТ ОИЯИ)

08.07.2021, 15:45

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The Multifunctional Information and Computing Complex in the Laboratory of Information Technologies of the Joint Institute for Nuclear Research is a multicomponent hardware and software complex, which ensures the fulfillment of a wide range of tasks related to the processing, analysis and storage of data in research conducted at the world level at JINR and in the world centers collaborating...

65. Complete decentralization of distributed data storages based on blockchain technology

Andrey Demichev (SINP MSU)

08.07.2021, 16:00

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The report presents a solution for completely decentralized data management systems in geographically distributed environments with administratively unrelated or loosely related user groups and in conditions of partial or complete lack of trust between them. The solution is based on the integration of blockchain technology, smart contracts and provenance metadata driven data management....

58. Modeling network data traffic for vulnerability scan using the TrafficREWIND test bench infrastructure of TIER1 data centers at JINR

Andrey Baginyan (ccnp)

08.07.2021, 16:00

2. Research infrastructure

Sectional reports

Research infrastructure

Modeling network data traffic is the most important task in the design and construction of new network centers and campus networks. The results of the analysis of models can be applied in the reorganization of existing

centers and in the configuration of data routing protocols based on the use of links. The paper shows how constant monitoring of the main directions of data transfer allows...

52. OpenMP computing of a reference solution for coupled Lorenz system on [0,400]

Dr Ivan Hristov (Sofia University, FMI, Bulgaria)

08.07.2021, 16:00

5. High Performance Computing

Sectional reports

HPC

Obtaining a long term reference trajectory on the chaotic attractor for coupled Lorenz system is a difficult task due to the sensitive dependence on the initial conditions. Using the standard double-precision floating point arithmetic, we cannot obtain a reference solution longer than 2.5 time units. Combining OpenMP parallel technology together with GMP library (GNU multiple precision...

70. Применение мультиагентных систем при обработке видеоданных

Mr Анатолий Каляев (Научно-исследовательского института многопроцессорных вычислительных систем имени академика А.В. Каляева Южного федерального университета)

08.07.2021, 16:00

1. Distributed computing systems

Distributed computing systems

В последние годы все большее распространение получают интеллектуальные системы видеообзора: охранные системы, системы анализа дорожной обстановки, системы выявления девиантного поведения, непрерывно растет число видеокамер, увеличивается разрешение получаемых изображений, усложняются алгоритмы обработки. Все это приводит к непрерывному увеличению генерируемой информации, и соответствующее...

75. Data Knowledge Base current status and operation

Viktor Kotliar (IHEP)

08.07.2021, 16:15

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The Data Knowledge Base (DKB) project is aimed at knowledge acquisition and metadata integration, providing fast response for a variety of complicated queries, such as summary reports and monitoring tasks (aggregation queries) and multi-system join queries, which are not easy to implement in a timely manner and, obviously, are less efficient than a query to a single system with integrated...

73. IPv6 dual-stack deployment for the distributed computing center

Viktor Kotliar (IHEP)

08.07.2021, 16:15

2. Research infrastructure

Sectional reports

Research infrastructure

Computing Center of the Institute for High Energy Physics in Protvino provides computing and storage resources for various HEP experiments (Atlas, CMS, Alice, LHCb) and currently operates more than 150 working nodes with around 3000 cores and provides near 2PB of disk space. All resources are connected through two 10Gb/s links to LHCONE and other research networks. IHEP computing center has...

147. Overlapping Computation and Communication in Matrix-Matrix Multiplication Algorithm for Multiple GPUs

Yea Rem Choi (HSE)

08.07.2021, 16:15

5. High Performance Computing

Sectional reports

HPC

In this talk, we discuss the optimal strategy for parallel matrix-matrix multiplication algorithm that minimizes the time-to-solution by finding the best parameters of the algorithm for overlapping multiplications of separate tiles in each GPU and data transfers between GPUs. The new algorithm developed for multi-GPU nodes is discussed [1]. The correlation is analyzed between the optimal...

121. Распределенные отказоустойчивые вычисления с SBN-Python на реальном кейсе компании

Mr Дмитрий Терещенко

08.07.2021, 16:15

1. Distributed computing systems

Sectional reports

Distributed computing systems

Распределённые вычисления сегодня достаточно востребованы в задачах пакетной обработки данных, но текущие решения, которые позволяют в Python их использовать, либо слишком узкоспециализированные, либо не дают полной отказоустойчивости.

В рамках выпускной квалификационной работы был разработан высокоуровневый интерфейс на Python (далее **SBN-Python**) к новому C++ фреймворку *Subordination*,...

133. Services of computational neurobiology tasks, based on the distributed modular platform «Digital Laboratory» NRC «Kurchatov Institute»

Irina Enyagina (Kurchatov Institute)

08.07.2021, 16:30

2. Research infrastructure

Sectional reports

Research infrastructure

This report will present services for performing computational neurobiology tasks for working with experimental data from nuclear magnetic resonance imaging of the human brain. These Services are created as separate modules based on the "Digital Laboratory platform" NRC "Kurchatov Institute".
On the basis of the distributed modular platform «Digital Laboratory» at NRC "Kurchatov Institute"...

136. Verifiable application-level checkpoint and restart framework for parallel computing

Mr Ivan Gankevich (Saint Petersburg State University)

08.07.2021, 16:30

5. High Performance Computing

Sectional reports

HPC

Fault tolerance of parallel and distributed applications is one of the concerns that becomes topical for large computer clusters and large distributed systems. For a long time the common solution to this problem was checkpoint and restart mechanisms implemented on operating system level, however, they are inefficient for large systems and now application-level checkpoint and restart is...

31. Развитие и совершенствование способов оценки качества технических систем в процессе эксплуатации

Камиль Закирович Билятдинов (Национальный исследовательский университет ИТМО, Санкт-Петербург)

08.07.2021, 16:30

1. Distributed computing systems

Sectional reports

Distributed computing systems

Предлагаются рациональные инструменты для инновационного развития и совершенствования способов оценки качества технических систем на этапе эксплуатации. Инновации основаны на том, что предлагаются для применения совокупность взаимосвязанных моделей, методик и реализующих их программ для ЭВМ, которые позволят сократить затраты времени и ресурсов на оценку качества систем и (или) элементов...

172. Development of information systems for theoretical and applied tasks on the basis of the HybriLIT platform

Yuri Butenko (JINR)

08.07.2021, 16:45

5. High Performance Computing

Sectional reports

HPC

The report gives an overview of two information systems (IS) under development on the basis of the HybriLIT platform. The major goal of creating these ISs is to automate calculations, as well as to ensure data storage and analysis for different research groups.

The information system for radiobiological research provides tools for storing experimental data of different types, a software set...

158. Precomputation formal verification of HPC cluster applications using SOPN Petri nets

Oleg Iakushkin (Saint-Petersburg State University)

08.07.2021, 16:45

2. Research infrastructure

Sectional reports

Research infrastructure

In this work, we present a formal mathematical model and software library for modelling hardware components and software systems based on the SOPN Petri network and CSharp programming language.

A discrete stochastic model denoted as SOPN, is presented, which combines the qualities of coloured, hierarchical and generalized Petri nets. The model is a series of extensions over the necessary...

95. Transformer-based Model for the Semantic Parsing of Error Messages in Distributed Computing Systems in High Energy Physics

Dmitry Grin

08.07.2021, 17:00

2. Research infrastructure

Sectional reports

Research infrastructure

Large-scale computing centers supporting modern scientific experiments store and analyze vast amounts of data. A noticeable number of computing jobs executed within the complex distributed computing environments ends with errors of some kind, and the amount of error log data generated every day complicates manual analysis by human experts. Moreover, traditional methods such as specifying...

175. Интеллектуальный анализ данных для повышения эффективности использования высокопроизводительной гетерогенной вычислительной платформы HybriLIT

Ekaterina Polegaeva (Dubna University)

08.07.2021, 17:00

5. High Performance Computing

Sectional reports

HPC

Гетерогенная вычислительная платформа HybriLIT является частью многофункционального информационно-вычислительного комплекса Лаборатории информационных технологий им. М.Г. Мещерякова Объединенного института ядерных исследований. Был проведен анализ данных по использованию платформы HybriLIT: особое внимание уделено исследованию информации об используемых ресурсах при запуске задач...

29. The grid-characteristic method for applied dynamic problems

Dr Vasily Golubev (Moscow Institute of Physics and Technology)

08.07.2021, 17:15

5. High Performance Computing

Sectional reports

HPC

Due to the rapid development of high-performance computing systems, more and more complex and time-consuming computer simulations can be carried out. It opens new opportunities for scientists and engineers. A standard situation for scientific groups now is to have an own in-house research software, significantly optimized and adopted for a very narrow scientific problem. The main disadvantage...

33. Accounting and monitoring infrastructure for Distributed Computing in the ATLAS experiment

Mr Aleksandr Alekseev (Ivannikov Institute for System Programming of the RAS)

09.07.2021, 09:00

4. Distributed computing applications

Plenary reports

Plenary reports

The ATLAS experiment uses various tools to monitor and analyze the metadata of the main distributed computing applications. One of the tools is fully based on the unified monitoring infrastructure (UMA) provided by the CERN-IT Monit group. The UMA infrastructure uses modern and efficient open-source solutions such as Kafka, InfluxDB, ElasticSearch, Kibana and Grafana to collect, store and...

41. Information technologies based on DNA. Nanobioelectronics

Mr Victor Lakhno (Institute of Mathematical Problems of Biology RAS – the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences)

09.07.2021, 09:45

1. Distributed computing systems

Plenary reports

Plenary reports

DNA molecular is a clear example of data storage and biocomputing. Performing millions of operations simultaneously DNA – biocomputer allows the performance rate to increase exponentially. The limitation problem is that each stage of paralleled operations requires time measured hours or days. To overcome this problem can nanobioelectronics.
The central problem of nanobioelectronics is...

94. Features of HPC resources for HEP

Artem Petrosyan (JINR)

09.07.2021, 11:00

5. High Performance Computing

Sectional reports

HPC

In the wake of the success of the integration of the Titan supercomputer into the ATLAS computing infrastructure, the number of such projects began to increase. However, it turned out that it is extremely difficult to ensure efficient data processing on such types of resources without deep modernization of both applied software and middleware. This report discusses in detail the problems and...

117. Increasing the accuracy of the diagnosis of mental disorders based on heterogeneous distributed data

Alexander Degtyarev (Professor), Alexander Bogdanov (St.Petersburg State University)

09.07.2021, 11:00

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The medical field, and especially diagnosis, is still an extremely poorly formalized field. This is especially true in the study of diseases associated with changes and disorders in the activity of the brain. In order to improve the results of medical research in this area, various methods of analyzing the condition of patients are used. These include both instrumental methods (MRI, EEG) and...

191. Neural network approach to the problem of image segmentation for morphological studies at LRB JINR

Oksana Streltsova (Meshcheryakov Laboratory of Information Technologies, JINR)

09.07.2021, 11:00

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

The report will present the results on the development of the algorithmic block of the Information System (IS) for radiobiological studies, created within a joint project of MLIT and LRB JINR, in terms of solving the segmentation problem for morphological research to study the effect of ionizing radiation on biological objects. The problem of automating the morphological analysis of...

71. Quantitative and qualitative changes in the JINR cloud infrastructure

Nikolay Kutovskiy (JINR)

09.07.2021, 11:00

7. Virtualization

Sectional reports

Virtualization

High demand for the JINR cloud resources facilitated its sufficient growth. That triggered changes needed to be done to overcome problems encountered and to keep QoS for users: main part of computational resources was re-organized as pre-deployed worker nodes of HTCondor-based computing element to decrease the load on OpenNebula services during mass jobs submission, new SSD-based ceph pool for...

192. Algorithms for behavioral analysis of laboratory animals in radiobiological research at LRB JINR

Alexey Stadnik (Meshcheryakov Laboratory of Information Technologies, JINR), Dina Utina (LRB JINR)

09.07.2021, 11:15

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

---

112. Energy analysis of plasma physics algorithms

Igor Chernykh (Institute of Computational Mathematics and Mathematical Geophysics SB RAS)

09.07.2021, 11:15

5. High Performance Computing

Sectional reports

HPC

High-performance supercomputers became one of the biggest power consumption machines. The top supercomputer's power is about 30mW. Recent legislative trends in the carbon footprint area are affecting high-performance computing. In our work, we collect energy analysis from different kinds of Intel's server CPUs. We present the comparison of energy efficiency of our new Poissons's solver, which...

76. Error detection in data storage systems and distributed voting protocols

Alexei Uteshev (St.Petersburg State University)

09.07.2021, 11:15

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The problems of silent data corruption detection in the data storage systems (Reed-Solomon codes) and faulty share detection in the distributed voting protocols (Shamir scheme) are treated from a uniform point of view. Namely, the both can be interpreted as the problem of systematic error detection in the data set { (x_1, y_1),...(x_N,y_N)} generated by a polynomial function y=f(x) in some...

174. The use of distributed clouds for scientific computing

Igor Pelevanyuk (Joint Institute for Nuclear Research)

09.07.2021, 11:15

7. Virtualization

Sectional reports

Virtualization

Nowadays, cloud resources are the most flexible tool to provide access to infrastructures for establishing services and applications. But, it is also a valuable resource for scientific computing. In the Joint Institute for Nuclear Research computing cloud was integrated with the DIRAC system. That allowed submission of scientific computing tasks directly to the cloud. With that experience, the...

125. ON DEEP LEARNING FOR OPTION PRICING IN LOCAL VOLATILITY MODELS

Sergey Shorokhov (RUDN University)

09.07.2021, 11:30

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Existence of exact closed-form formula for the price of derivative is a rather rare event in derivative pricing, therefore, to determine the price of derivative, one has to apply various numerical methods, including finite difference methods, binomial trees and Monte Carlo simulations. Alternatively, derivative prices can be approximated with deep neural networks.

We study pricing of...

135. Passwordless Authentication Using Magic Link Technology

Iurii Matiushin (Saint Petersburg State University)

09.07.2021, 11:30

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

Nowadays, the problem of identification and authentication on the Internet is more urgent than ever. There are several reasons for this: on the one hand, there are many Internet services that keep records of users and differentiate their access rights to certain resources; on the other hand, cybercriminals' attacks on web services have become much more frequent lately. At the same time, in...

114. VM based Evaluation of the Scalable Parallel Minimum Spanning Tree Algorithm for PGAS Model

Vahag Bejanyan

09.07.2021, 11:30

5. High Performance Computing

Sectional reports

HPC

The minimum spanning tree problem has influential importance in computer science, network analysis, and engineering. However, the sequential algorithms become unable to process the given problem as the volume of the data representing graph instances overgrowing. Instead, the high-performance computational resources pursue to simulate large-scale graph instances in a distributed manner....

182. Минимизация образов корневых файловых систем Docker контейнеров

Irina Nikolaeva

09.07.2021, 11:30

7. Virtualization

Sectional reports

Virtualization

В последнее время при разработке и тестировании приложений достаточно часто обращаются к контейнерам. Это обусловлено тем, что контейнер – это удобный в использовании и легковесный инструмент. Контейнер собирается на основе образа, который является шаблоном будущего контейнера, а также с помощью образа контейнер передается по сети. Размер образов может достигать нескольких гигабайт. Таким...

180. Analytical platform for socio-economic studies

Sergey Belov (Joint Institute for Nuclear Research)

09.07.2021, 11:45

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Started in natural sciences, the high demand for analyzing a vast amount of complex data reached such research areas as economics and social sciences. Big Data methods and technologies provide new efficient tools for researches. In this paper, we discuss the main principles and architecture of the digital analytical platform aimed to support socio-economic applications. Integrating specific...

178. Data storage systems of "HybriLIT" Heterogeneous computing platform for scientific research carried out in JINR: filesystems and RAIDs performance research

Aleksander Kokorev

09.07.2021, 11:45

5. High Performance Computing

Sectional reports

HPC

"HybriLIT" Heterogeneous platform is a part of the Multifunctional Information and Computing Complex (MICC) of the Laboratory of Information Technologies named after MG Meshcheryakov of JINR, Dubna. Heterogeneous platform consists of Govorun supercomputer and HybriLIT education and testing polygon. Data storage and processing system is one of the platform components. It is implemented using...

84. Resource Management in Private Multi-Service Cloud Environments

Mr Nikita Balashov (JINR)

09.07.2021, 11:45

7. Virtualization

Sectional reports

Virtualization

The JINR cloud infrastructure hosts a number of cloud services to facilitate scientific workflows of individual researchers and research groups. While the batch processing systems are still the major compute power consumers of the cloud, new auxiliary cloud services and tools are being adopted by researchers and are gradually changing the landscape of the cloud environment. While such...

151. RISK MODEL OF APPLICATION OF LIFTING METHODS

Mr Aleksandr Dik (Saint Petersburg State University)

09.07.2021, 11:45

6. Data Management, Organisation and Access

Sectional reports

Data Management, Organization and Access

The article discusses the main provisions (methods, risk models, calculation algorithms, etc.) of the issue of organizing the protection of personal data (PD), based on the application of anonymization procedure. The authors reveal the relevance of the studied problem based on the tendency of the general growth of informatization and the further development of the Big Data technology. This...

127. Evaluating Different Options for Scientific Computing in Public Clouds

Dr Oleg Sukhoroslov (IITP RAS, NRU HSE)

09.07.2021, 12:00

7. Virtualization

Sectional reports

Virtualization

Cloud computing has emerged as a new paradigm for on-demand access to a wast pool of computing resources that provides a promising alternative to traditional on-premises resources. There are several advantages of using clouds for scientific computing. Clouds can significantly lower time-to-solution via quick resource provision, skipping the lengthy process of building a new cluster on-premises...

140. HPC workload balancing algorithm for co-scheduling environments.

Ruslan Kuchumov (Saint Petersburg State University)

09.07.2021, 12:00

5. High Performance Computing

Sectional reports

HPC

Commonly used job schedulers in high-performance computing environments do not allow resource oversubscription. They usually work by assigning the entire node to a single job even if the job utilises only a small portion of nodes' available resources. This may lead to cluster resources under-utilization and to an increase of job wait time in the queue. Every job may have different requirements...

118. Solving the problems of Byzantine generals using blockchain technology

Jasur Kiyamov (St.Petersburg State University)

09.07.2021, 12:00

6. Data Management, Organisation and Access

Data Management, Organization and Access

The process of digitalization of the Russian economy as the basis for the transition to the digital economy is conditioned by the requirements of objective reality and is based, first of all, on the introduction of digital technologies into the activities of its actors. The most promising is the Blockchain technology, which has the capabilities of the most effective coordination of the...

137. Применение методов машинного обучения для кросс-классификации алгоритмов и задач многомерной непрерывной оптимизации

Mr Андрей Чепурнов (ф-т ВМК МГУ им. М.В. Ломоносова)

09.07.2021, 12:00

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

Предлагаемая работа посвящена разработке программной системы для проведения взаимной классификации семейств популяционных алгоритмов оптимизации и задач многомерной непрерывной оптимизации. Одной из целей настоящего исследования является разработка методов предсказания эффективности работы включенных в систему алгоритмов и выбора из них наиболее эффективных алгоритмов для решения заданной...

163. Characteristics of Nvidia CUDA and AMD ROCm Platforms Affecting Performance Portability

Vsevolod Nikolskiy (HSE)

09.07.2021, 12:15

5. High Performance Computing

Sectional reports

HPC

The development and popularization of the AMD ROCm platform with HIP technology allows one to create code that is not locked to a specific vendor maintaining a high level of performance. A lot of legacy but still supported codes is originally written in CUDA, and now it is getting ROCm HIP support as well. In a recent paper [1], the performance of popular molecular dynamics packages with GPU...

155. Multi-instance learning for Rhetoric Structure Parsing

Mr Sergey Volkov (Peoples' Friendship University of Russia (RUDN University); Federal Research Center "Computer Science and Control" RAS)

09.07.2021, 12:15

9. Big data Analytics and Machine learning

Sectional reports

Big data Analytics and Machine learning.

To accurately detect texts containing elements of hatred or enmity, it is necessary to take into account various features: syntax, semantics and discourse relations between text fragments. Unfortunately, at present, methods for identifying discourse relations in the texts of social networks are poorly developed. The paper considers the issue of classification of discourse relations between two...

63. Применение новых технологий виртуализации в учебном процессе

David Satseradze (Professor, Lecture)

09.07.2021, 12:15

7. Virtualization

Sectional reports

Virtualization

Современное развитие IT-инфраструктуры осуществляется форсировано во время которого происходит переход на облачные вычисления с внедрение методов виртуализации.
Наряду с этим, в настоящее время, эффективное решение крупномасштабных научных задач требует применения высокопроизводительных вычислений, в том числе использования распределенных вычислительных сред (РВС) различного назначения.
Все...

216. Advantages of overlay scientific networks creation on top of RETN core optical transport network

Andrey Ul'yanov (RETN)

Research infrastructure

193. STATISTICAL MECHANICS APPROACH FOR DEEP-BELIEF NEURAL NETWORKS EXPLORATION

R. Rudamenko (Faculty of Physics, Lomonosov Moscow State University)

Sectional reports

Distributed computing, HPC and ML for solving applied tasks

Humans and other animals can understand concepts from only a few examples. while standard machine learning algorithms require a large number of examples to extract hidden features. Unsupervised learning is procedure of revealing hidden features from unlabeled data.
In deep neural network training, unsupervised data pre-training increases the final accuracy of the algorithm by decreasing an...

211. To be announced later